There's More Than One Way to Skin a CAT

CAT Tools

Computer Assisted Translation tools: all tools that aid in the translation process

- Project management applications
- Term extraction tools
  - *Xerox XTS*, *LogiTerm*, *TermExtract*, *SDLPhraseFinder*
- Word count tools
  - *PractiCount*, *AnyCount*
- Localization tools
  - *Catalyst*, *Passolo*, *RC-Wintrans*
- Machine translation programs

Project Management/Workflow Tools

Translation process orientation

Business process orientation

The Translation Process Tools

- *TRADOS TeamWorks*

- *Star Translation Workflow Server*

- *Idiom WorldServer*

The Business Process Tools

- *LTC Organiser* (www.langtech.co.uk)

- *]project-open[* (www.project-open.com)

- *Projetex/Translation Office 3000* (www.to3000.com)

- *T.O.M.* (www.jovo-soft.de)

Workflow Tools: The Questions to Ask

- Online interface
  - hosted?
- API or interface to accounting and/or TM software
- Size of existing implementation
  - capacity of database
- Background of developer
- Customizability

Can I change for the system or can the system change for me?

"Disdain on the side of the professional translators for the hilarious and stupid MT mistakes gave birth to a new variant of MT called translation memory (TM). TM started off as a lower-level feature of commercial MT systems (...). But the success of TM came with dedicated products such as IBM TM/2 and Trados. The marketing message was tuned in to what the professional translation industry wanted to hear: 'Forget about MT; it doesn't work well. Instead, use our TM product because it leaves you in full control of the process.'

"The message worked well. Within a period of 10 to 15 years, TM products have found their way to the workstations of more than 50,000 translators in the world. But the message has also caused a 'cognitive disorder' in the translation industry, namely that TM is good and MT is evil, foregoing the fact that TM is just a new variant of MT (...). The damage is done, however, and it will take years to convince the community of business translators that post-editing fuzzy matches from TM databases is, in fact, not different from post-editing fuzzy matches from any other MT system."

Jaap van der Meer in *MultiLingual* 71 (2005)

Categories of Machine Translation

- Linguistic rule-based

- Data-driven or EBMT/SBMT

- Hybrid approach

Defining Translation Memory

"Applications that extend the memory of the translation professional by allowing him or her to build up databases of translated material and leverage that against newly translatable content."

In reality they're usually more than that because most of them

- allow the user to build up and maintain terminology databases

- allow translators and project managers to work in complex file formats

- provide methods for analysis

- provide quality assurance

- provide project management capabilities

Categories of TM Tools

- Tools that perform all or most of their work through macros in *Microsoft Word*

- Tools that let the translator work in an independent environment

Categories of TM Tools

Tools that use Microsoft Word as their main translation interface

Trados (www.trados.com)

MultiTrans (www.multicorpora.com)

Wordfast (www.wordfast.net)

MetaTexis (www.metatexis.com)

WordFisher (www.wordfisher.com)

Fusion (www.orcadev.com)

Categories of TM Tools

Tools that use an independent translation interface

ForeignDesk (www.foreigndesk.net)

across (www.across.net)

Heartsome (www.heartsome.net)

OmegaT (www.omegat.net)

SDLX (www.sdlx.com)

Star Transit (www.star-transit.com)

Déjà Vu (www.atril.com)

Translation Memory

Similarities of the major tools (*Trados*, *SDLX*, *Transit*, *Déjà Vu*):

- Freely configurable terminology component
- Unicode and TMX support
- Alignment feature
- Similar range of supported file formats
- Concordance searches
- Analysis features

Translation Memory

Things to look for:

- Code handling

  HTML: <FONT face=Arial size=2><P>The <B>cat</B> is black.</P></FONT>

  Trados: `<FONT face=Arial size=2>` `<P>`The `<B>`cat `</B>` is black.`</P>` `</FONT>`

  Transit: <F>>«3+»The **cat** is black.«4»<<F>

  Déjà Vu: The {29}cat{30} is black.

  SDLX: The **cat** is black.

- Supported file formats

- Batch vs. individual file processing

- Support

Terminology Handling

The Prejudices:

- "Translation memory" program seems to suggest that the emphasis is on the translation memory
- More immediate gain through perfect and fuzzy matches on a sentence-by-sentence basis
- Translation memories can be relatively quickly built up by aligning existing translated file pairs and/or automatically by translating new texts
- The construction of terminology databases is a comparatively tedious process

Terminology Handling

The Benefits:

- The terminology database is the place to invest effort into defining words and phrases grammatically, contextually, or even by contrast
- "Living dictionaries" that present their findings for each segment being translated
- Term entries contain meta data that intelligently determine relevance
- Consistency throughout workgroups

Terminology Handling

The Highlights:

- Web-based versions for *Transit*, *SDLX* and *Trados*
- "Assembling" segments that are missing from TM and "fixing" fuzzy matches in *Déjà Vu X*
- Revamped entry method in *Trados*
- TBX

The purchase of a software tool is only the beginning of the investment into translation memory technology.

Without training the investment will fail.

How to Fail

"In my opinion, any application interface must be intuitive. If it isn't, if an application requires any additional training, it's a poorly designed application. TRADOS is one example of such applications. On the other hand, applications made by Microsoft are examples of good design. Yes, they make money on training, however they do not teach people how to use MS Word because it is not necessary."

A Glance into the Future

Content Management Systems:

Trados/SDL has a number of partnerships with content management providers and provides a increasingly comprehensive authoring/translation environment

Challenge to the language provider:

– With a true integration of content management and translation memory, either the clients will be more heavily involved in translation memory management or the vendors will have to broaden their service portfolio.

A Glance into the Future

Workflow/Project Management:

- Increasingly strong emphasis *on* workflow component in TM tools.

- Most language vendors today would describe the need of an adequate workflow system as equally if not more important than a TM system.

A Glance into the Future

Online Access:

- Most newer translation memory systems (*Fusion*, *Logoport*, *across*) have online access as the core component. More well-known tools (*Trados*, *Transit*, and *SDLX*) also have at least one online access component.

  Virtual workgroups and routine high-speed access have long become a reality. The only effective way to exchange data is the use of common databases.

A Glance into the Future

TM as a Marketable Asset:

In the past the discussion surrounding this did not progress beyond the copyright and quality concerns.

Interest in a marketplace for translation memory data has long been present (and a reality as expressed in the popularity of the Microsoft "glossaries").

End clients with copyrights to their translation memories can receive a more immediate return on investment of their translation costs.

A Glance into the Future

Official Reunification of TM and MT Technology:

Translation vendors and buyers will understand limitations and challenges

Data is more available than ever before

Representatives of the MT community will be on the ATA board

The Translator's Tool Kit

www.internationalwriters.com/toolkit